

中图法分类号: 文献标识码: 文章编号: 1006-8961(XXXX)XX-0001-17

论文引用格式: BAO Xiao-an, Chen Yi-jiang, ZHANG Na, Hu Tian-bin, TU Xiao-mei. XXXX. Zero-Supervised Low-Light Image Enhancement with Illumination-Guided Reflectance Estimation. Journal of Image and Graphics, XX(XX):0001-0017(包晓安, 陈奕江, 张娜, 胡天缤, 许铭洋, 涂小妹. XXXX. 光照引导反射分量估计的零监督弱光图像增强. 中国图象图形学报, XX(XX):0001-0017)[DOI:10.11834/jig.250486]

光照引导反射分量估计的零监督弱光图像增强

包晓安¹, 陈奕江¹, 张娜¹, 胡天缤², 许铭洋¹, 涂小妹³

1. 浙江理工大学计算机科学与技术学院, 浙江杭州 310018; 2. 河海大学人工智能与自动化学院, 江苏常州, 213000; 3. 浙江广厦建设职业技术大学建筑工程学院, 浙江东阳 322100

摘要: 目的 弱光环境下成像质量下降严重制约视觉监控、自动驾驶等高级视觉任务的性能。现有方法虽可提升亮度,但普遍存在过曝抑制不足且依赖成对正常光照数据的问题,限制了其在真实场景中的应用。方法 针对这些局限,提出一种基于光照引导反射分量估计的零监督弱光图像增强方法。首先,通过可逆亮度扰动合成训练样本,设计伪曝光生成策略,在不改变图像固有反射率的情况下,生成多种光照样本以摆脱成对监督数据的依赖;其次,设计双域协同注意力去噪模块以抑制背景噪声;然后,利用多层残差卷积与结构对称的U-Net架构协同生成照明图,并利用其深层光照信息对反射分量进行显式引导;最后,通过引导式反射估计实现结构纹理的高保真重建与过曝抑制。结果 在LOL、DICM和WildNight-8K数据集上与9种代表性的最新方法进行对比分析。在LOL数据集上,相比Retinexformer和LightenDiffusion,本文方法的自然图像质量评估指标(natural image quality evaluator, NIQE)分别降低0.62和0.92,感知指数(perceptual index, PI)分别降低1.23和0.25;在DICM数据集上,相比Retinexformer和RUAS,本文方法的NIQE分别降低0.11和0.81,PI分别降低0.48和0.72;在WildNight-8K数据集上,相比Retinexformer和EnlightenGAN,本文方法的NIQE分别降低1.32和1.87,PI分别降低0.33和5.98。在YOLOv8目标检测任务中,平均检测精度(mean average precision, mAP@50-95)较原图提升4.60%。结论 本文提出的基于光照引导反射分量估计的零监督弱光图像增强方法在多个数据集上取得了领先的客观指标,并能够有效提升下游目标检测任务性能,验证了其在真实弱光应用场景中的价值。

关键词: 弱光图像增强;注意力机制;零监督;目标检测;卷积神经网络

Zero-Supervised Low-Light Image Enhancement with Illumination-Guided Reflectance Estimation

BAO Xiao-an¹, Chen Yi-jiang¹, ZHANG Na¹, Hu Tian-bin², TU Xiao-mei¹

1. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China; 2. School of Artificial Intelligence and Automation, Hohai University, Changzhou 213000, China; 3. School of Civil Engineering and Architecture, Zhejiang Guangsha Vocational and Technical University of Construction, Dongyang 322100, China

Abstract: Objective Low-light environments in practical computer vision applications, such as security surveillance, autonomous driving, and intelligent urban monitoring, significantly degrade image quality and thus restrict the performance of high-level visual tasks. Images captured under poor illumination often exhibit insufficient brightness, heavy noise ampli-

收稿日期: 2025-10-02; 修回日期: 2026-01-28

基金项目: 浙江省重点研发计划项目(2020C03094); 金华市科技局重点课题(2024C22592)

Supported by: Zhejiang Provincial Key Research and Development Program(2020C03094); Jinhua Municipal Science and Technology Bureau Key Research Project(2024C22592)

fication, and severe loss of structural and textural information. These degradations not only harm perceptual quality but also compromise the reliability of downstream applications such as object tracking, object detection, and semantic segmentation. Existing methods have attempted to solve these problems by increasing image brightness, but they frequently suffer from inadequate overexposure suppression and heavy reliance on paired datasets collected under normal lighting conditions. Since such datasets are expensive and difficult to obtain, the generalization ability of these supervised methods remains limited in real-world complex scenarios. Furthermore, some unsupervised methods attempt to reduce dependency on paired data, but they often fail to simultaneously achieve noise suppression, exposure correction, and structure preservation, leading to unsatisfactory performance in practice. **Method** To address these challenges, this study proposes a novel zero-supervision low-light image enhancement framework that utilizes illumination components to guide reflectance estimation, thereby ensuring high-fidelity reconstruction. The framework is designed to operate without paired training data, while simultaneously achieving effective noise suppression, overexposure control, and structure preservation. The process begins with a reversible brightness perturbation strategy combined with pseudo-exposure generation, which synthesizes abundant pseudo low-light and pseudo well-lit samples without changing the intrinsic reflectance of the image. This effectively simulates diverse illumination conditions and eliminates the dependence on paired supervision. To further improve robustness, a dual-domain collaborative attention denoising module is introduced. This module operates in both the spatial domain and the channel domain: in the spatial domain, contextual dependencies are captured to distinguish true structures from random noise, while in the channel domain, correlations among different feature channels are modeled to suppress redundant information and emphasize critical features. By collaborating across these two domains, the module ensures effective denoising while preserving structural fidelity. For structural representation, multi-layer residual convolutions and a structurally symmetric U-Net architecture work in concert to generate the illumination map. The symmetric encoder-decoder structure enables extraction of multi-scale features, capturing both global illumination and local detail, while the residual connections mitigate vanishing gradients and accelerate convergence. The U-Net is designed not only to generate illumination maps but also to extract illumination-related features that serve as explicit guidance for reflectance recovery. Based on this, a guided reflectance estimation method is introduced. Instead of directly treating reflectance as a residual, illumination priors are fused into the estimation process. This enables the network to effectively separate illumination variations from intrinsic structures, restoring edges and textures with high fidelity while suppressing overexposed regions. The integration of illumination priors makes the restored images appear both natural and structurally consistent, thereby addressing the shortcomings of previous enhancement methods. **Result** Extensive comparative experiments were conducted on the LOL, DICM, and WildNight-8K datasets against nine representative state-of-the-art low-light image enhancement methods. On the LOL dataset, the proposed method outperformed Retinexformer and LightenDiffusion, achieving reductions of 0.62 and 0.92 in the Natural Image Quality Evaluator (NIQE), as well as decreases of 1.23 and 0.25 in the Perceptual Index (PI), respectively. On the DICM dataset, compared with Retinexformer and RUAS, the proposed approach reduced NIQE by 0.11 and 0.81, and PI by 0.48 and 0.72, respectively. On the WildNight-8K dataset, relative to Retinexformer and EnlightenGAN, NIQE was further reduced by 1.32 and 1.87, while PI decreased by 0.33 and 5.98, respectively. These results demonstrate that the proposed method consistently delivers enhanced images with improved naturalness and perceptual quality across diverse low-light scenarios. Beyond image quality assessment, the enhanced images were further evaluated in a downstream animal detection task using YOLOv8. Experimental results show that the mean average precision (mAP@50-95) increased by 4.60% compared with the original low-light images, indicating that the proposed enhancement method not only improves visual quality but also significantly boosts the reliability of downstream perception tasks in challenging low-light conditions. **Conclusion** This study introduces a zero-supervision low-light image enhancement framework that integrates illumination-guided reflectance estimation for high-fidelity image reconstruction. The key innovations include reversible brightness perturbation with pseudo-exposure generation to remove paired dataset dependency, a dual-domain collaborative attention denoising module to reduce noise while preserving structural details, a residual U-Net backbone for effective multi-scale representation, and an illumination-fused reflectance estimation mechanism to achieve structure-preserving restoration and overexposure suppression. Experimental results across multiple datasets and real-world applications demonstrate that the proposed method not only achieves state-of-the-art performance in quan-

titative and qualitative metrics but also significantly improves downstream vision tasks, particularly object detection. These contributions indicate strong practical value and broad application potential. Looking forward, future work could explore extending this framework to dynamic video enhancement with temporal consistency, incorporating adaptive illumination priors derived from scene semantics, and optimizing lightweight deployment for edge devices such as autonomous vehicles and mobile cameras. With these directions, the proposed method could further strengthen its role as a reliable solution for low-light visual perception in safety-critical and real-time applications.

Key words: low-light image enhancement; attention module; zero-reference; object detection; convolutional neural network

在此处键入公式。中图法分类号: TP391 文献标识码: A 文章编号: 论文引用格式:

0 引言

弱光图像增强作为计算机视觉与图像处理中的基础问题,对于提升图像视觉质量并保障下游感知任务的可靠性具有重要意义(Liu等,2023)。低照度成像常伴随亮度不足、补光过曝、噪声放大与细节缺失等问题,严重影响了图像的可用性和可解释性(Chen等,2024)。

针对这一问题,研究者提出了多类方法。早期传统方法主要基于直方图均衡、自适应直方图均衡及对比度拉伸等简单增强技术,通过亮度与局部对比度调整改善成像效果。然而在复杂弱光场景下,这类方法往往难以有效抑制噪声,并可能造成纹理细节的进一步丢失,表现出较强的局限性(Guo等,2017)。

近年来,深度学习驱动的弱光增强方法取得了快速发展(Xu等,2023)。以卷积神经网络(convolutional neural networks, CNN)为代表的早期工作,多基于曲线估计或图像分解,学习从低光到正常光的端到端映射(Dong等,2022)。然而,CNN方法在处理光照高度不均匀的自然场景时常面临结构保持不足与泛化性能受限的问题。随后,Retinex理论与深度模型的结合成为重要发展方向:如Retinexformer(Cai等,2023)通过引入变换器结构提升长程依赖建模与非均匀光照下的结构保持能力,RetinexMamba(Bai等,2024)则利用状态空间模型(state space model, SSM)在较低计算成本下获得近乎Transformer的表现。同时,CIDNet(Yan等,2025)提出可训练HVI色彩空间,该空间中引入可学习映射来替代固定的RGB表征,从而在弱光图像增强过程中更

好地分离亮度与色彩信息,这一设计能够有效缓解常见的色偏与色彩不稳定问题,尤其是在长时间曝光或复杂光照条件下,提升了增强结果的自然感与一致性(Sun等,2025)。FLOL(Benito等,2025)在频域与空域轻量融合的框架下实现了毫秒级推理,展现了在实际部署中的应用潜力。总体而言,Transformer与SSM的引入显著提升了表征与光照建模能力,而色彩空间建模与轻量化路线改善了效率与稳定性,弥补了传统CNN在真实夜景中的不足。

与此同时,扩散模型的兴起为弱光增强带来了新的思路。去噪扩散概率模型(denoising diffusion probabilistic model, DDPM)(Ho等,2020)最初应用于图像生成,在弱光图像中被引入为从退化的低光图像逐步恢复清晰、高质量图像的途径。Somepalli等人提出的DiffusionRetinex方法(Somepalli等,2023)利用扩散过程建模图像在不同曝光条件下的潜在分布,在去噪的同时增强了亮度与细节信息。Jiang等人提出的LightenDiffusion方法(Jiang等,2024)将Retinex分解迁移至潜空间,并采用扩散模型进行无监督恢复,同时引入自约束一致性损失以抑制内容泄漏。在多个真实场景基准数据集上的结果显示,该方法相较于同类无监督方法表现更优,并接近有监督方法的水平,同时具备更好的场景泛化能力。但是,与Transformer相比,基于扩散的增强方法在推理时延上仍然存在劣势,通常需要额外的加速或蒸馏策略来满足实际应用需求(Lin等,2025)。从近三年(2023-2025)的研究进展来看,弱光图像增强方法在建模范式与应用目标上呈现出明显分化趋势。以Retinexformer(Cai等,2023)和RetinexMamba(Bai等,2024)为代表的方法,通过引入Transformer或状态空间模型(SSM)显著增强了全局光照建模能力,在非均匀照明和大尺度亮度变化场景下取得了优于传统CNN的效果。然而,这类方法通常依赖复

杂的网络结构与多阶段特征交互,对训练数据分布较为敏感,在零监督或真实复杂环境中仍存在一定泛化风险。在此基础上,部分研究开始关注弱光增强中的色彩稳定性与感知一致性问题,CIDNet(Yan等,2025)通过构建可训练的HVI色彩空间,实现了亮度与色彩信息的更有效解耦,在长曝光和复杂光照条件下显著缓解了色偏现象。但该类方法主要从颜色表征层面改进增强效果,对噪声放大与结构退化问题的抑制能力相对有限,往往需要与额外的去噪或结构约束模块协同使用。另一方面,扩散模型为弱光增强提供了更强的分布建模能力。Diffusion-Retinex(Somepalli等,2023)与LightenDiffusion(Jiang等,2024)通过逐步去噪的生成过程,在细节恢复和噪声抑制方面展现出优势,尤其在无监督条件下具备较好的泛化性能。然而,扩散模型普遍存在推理开销大、亮度调节效率低以及局部过曝等问题,限制了其在实时或轻量化应用场景中的实用性。此外,FLOL(Benito等,2025)等最新的零监督轻量化方法通过频域与空域的高效融合显著提升了推理速度,但在极端弱光和高噪声条件下仍难以兼顾增强质量与结构保真度。综合来看,近三年的方法在全局建

模、色彩一致性或生成质量等单一维度上均取得了进展,但在零监督条件下同时兼顾噪声抑制、结构保持、过曝控制与计算效率方面仍存在不足。这也进一步凸显了在无需成对数据的前提下,构建具备稳定分解能力与鲁棒增强性能的弱光图像增强框架的研究价值。

在上述对比分析的基础上可以看出,国内外弱光图像增强的研究已从传统直方图与对比度增强,发展到CNN驱动的曲线估计与分解网络,再到融合Retinex先验的Transformer/SSM框架,以及近年的扩散模型(Li等,2025)。它们在亮度提升、噪声抑制与结构保持等方面不断取得突破,但整体仍面临三大挑战:其一,非均匀光照与补光噪声导致增强结果易出现局部过曝与失真;其二,复杂背景下缺乏有效结构建模,增强过程容易破坏目标边界与纹理细节(Wei等,2018),从而影响下游检测与分割性能;其三,成对监督数据采集成本高昂,限制了现有方法在零监督或弱监督条件下的适用性(Lin等,2025)。基于此,亟需一种无需成对数据、能够在真实复杂环境下实现弱光增强,同时兼顾噪声抑制、结构保真与视觉真实性的新型方法。

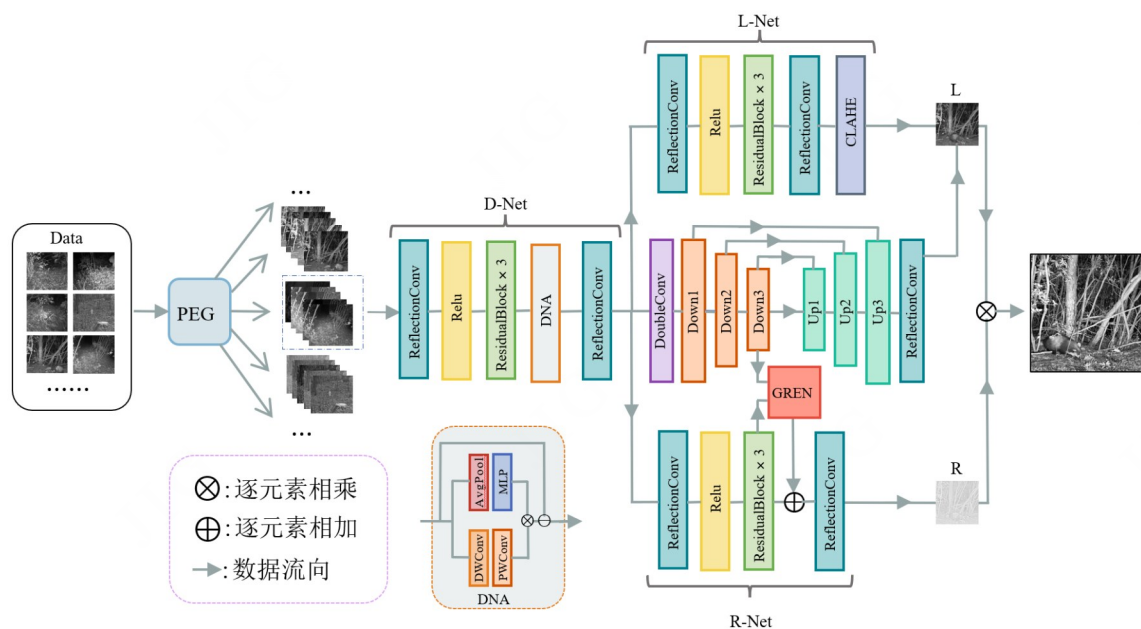


图1 本文网络结构

Fig. 1 The network structure of this article

1 本文方法

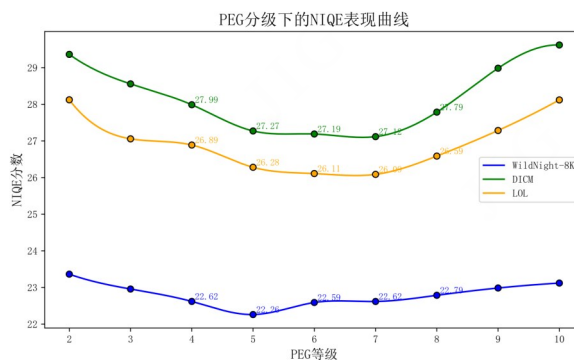
本文提出的零监督弱光增强框架如图 1 所示, 首先, 低光图像首先由伪曝光生成策略 (Pseudo-Exposure Generation, PEG) 处理, 通过对输入施加可逆亮度变换生成多幅伪曝光图, 为后续分解提供无标签训练样本。随后, 图像进入带有噪声抑制模块 (Denoise Attention, DNA) 的 D-Net 完成自适应去噪, 以获得干净的输入特征。接着 L-Net 和 U-Net 结构估计光照图 L ; R-Net 在融合 U-Net 提取的深层照明特征的基础上, 通过引导式反射图增强模块 (Guided Reflectance Enhancement Network, GREN) 重建反射图 R 。最终将光照与反射逐像素相乘, 输出亮度自然且纹理保真的增强结果。

1.1 可逆亮度扰动的伪曝光生成

在 Retinex 理论框架下, 图像 I 可被建模为光照分量 L 与反射分量 R 的乘积, 即: $I = R \odot L$ (Land 等, 1971)。其中, R 表征场景固有的材质与纹理属性, L 则反映入射光照的分布。传统的监督式 Retinex 分解方法通常依赖具有严格曝光配对关系的成对图像与正常光图, 以此提供对 L 和 R 的监督信息 (Guo 等, 2020)。然而, 高质量成对曝光数据在真实拍摄条件下极难获取, 严重限制了此类方法在实际场景中的应用范围。为突破成对标签依赖的瓶颈, 本文采用了一种零监督的分解策略。具体而言针对单张低光输入图像 I , 本文通过伪曝光生成策略 (PEG) 施加可逆的亮度变换, 采用 γ 校正、线性缩放、对数/指数映射, 从而生成多个亮度水平不同的伪曝光图像 $\{I_j\}$ 。在 Retinex 假设下, 这些伪图共享相同的反射信息 R , 但其光照分量 L_j 存在差异。因此, 即便缺乏真实曝光标签, 也可将它们构造成训练样本, 驱动网络学习分离出一致的 R 与各自对应的 L_j 。训练目标通过重建损失、一致性约束、平滑正则等多项损失函数共同优化, 逐步引导网络从非成对图像中学习到结构一致、物理意义明确的光照与反射分离。这种策略摆脱了对昂贵成对数据的依赖, 在保持训练稳定性的同时显著提升了模型的实际应用能力。图 2 PEG 分级下的 NIQE 表现曲线

Fig. 2 NIQE performance across PEG levels

需要指出的是, 真实相机曝光变化本质上由入射光通量、曝光时间及传感器响应函数等因素共同



决定, 其对图像亮度分布的影响通常表现为近似的非线性映射关系。在未知具体相机参数的情况下, γ 校正、线性强度缩放以及对数/指数映射等操作能够在统计意义上有效近似不同曝光条件下的亮度变化过程 (Wang 等, 2023)。在此基础上, 本文所提出的伪曝光生成策略 (PEG) 并非试图精确复现某一具体相机的物理曝光模型, 而是从数据驱动与 Retinex 分解一致性的角度出发, 通过多种可逆亮度变换构造一组覆盖不同亮度分布的伪曝光样本, 以最大化光照分量的多样性。由于这些变换对反射分量的影响可视为一致, 从而满足 Retinex 假设中“反射不随光照变化”的基本前提。

表 1 曝光变化方法对比分析表

Table 1 Comparative Analysis Table of Exposure Variation

曝光变化方法	NIQE↓	DRN↓	PI↓
单一 γ 校正	23.5027	0.2278	17.5984
线性增益调整	24.1502	0.2401	17.4502
基于直方图重映射	24.8578	0.2541	19.2481
PEG	22.261	0.1814	15.5665

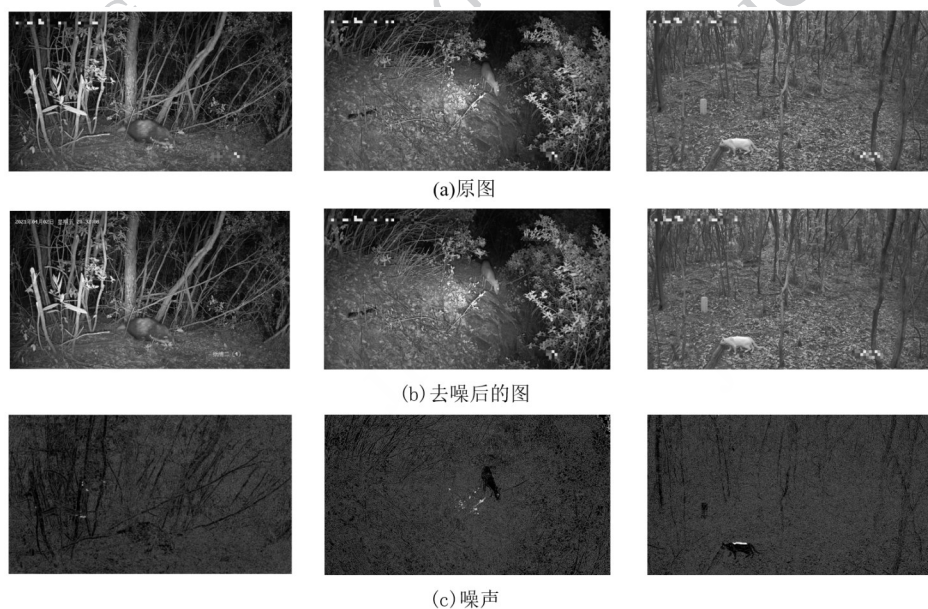
注: 加粗字体表示最优结果, 箭头向下表示负向指标, 数值越低代表表现越优。

为进一步验证 PEG 在模拟曝光变化方面的有效性, 本文将其与多种常见曝光变换策略进行了对比分析, 包括单一 γ 校正、线性增益调整和基于直方图重映射的亮度增强方法。不同策略生成的伪曝光图像分别用于 Retinex 分解网络训练, 结果如表 1 所示, 相较于单一形式的曝光变换, PEG 所生成的多级伪曝光在 NIQE 无参考质量指标上取得了更优的表现, 说明其在提升光照建模多样性与训练稳定性方面具有明显优势。

同时,PEG数量的选择直接影响生成伪曝光图像的覆盖范围及模型的泛化能力。当PEG数量过少时,生成的伪曝光样本亮度分布有限,光照分量多样性不足,网络在分离 R 与 L 时容易陷入欠拟合,导致增强后的低光图像仍保留暗区噪声或亮度不均匀;相反,当PEG数量过多时,虽然理论上可以覆盖更广亮度范围,但过多的伪曝光图像会引入大量冗余样本,增加网络训练负担,降低收敛速度,且在极端低光、结构复杂且噪声繁多的图像中可能出现伪影,从而降低图像自然度和NIQE指标表现。为验证PEG数量对不同光照条件下增强效果的影响,本文进行了多级别PEG数量的实验,并通过量化指标对各级别的增强效果进行评估,结果如图2所示。对于极低光照、噪声多、结构复杂的场景(如Wild-Night-8K数据集),PEG数量设置为5时,网络能够在保证光照分量多样性的同时避免冗余样本带来的训

练负担,因此在NIQE指标上表现最佳,这表明适度的伪曝光级别可在极端低光条件下最大化模型的泛化能力与图像自然度。相比之下,对于亮度稍高、噪声较少且结构相对简单的场景,增加PEG数量至7可以进一步扩展训练样本覆盖范围,使网络在更宽亮度分布下学习到更稳健的光照表示,从而获得略优的增强效果。

这一趋势表明,PEG数量的最优设置具有一定的条件依赖性:极暗且噪声多的图像倾向于较少的PEG级别,而亮度较好、场景简单的图像则可适当增加PEG数量以获取更丰富的光照多样性。在本研究中,PEG=5可作为低光增强任务的默认优选值,兼顾训练效率与泛化性能;同时,对于特定光照条件或不同场景复杂度的图像,可根据亮度分布与噪声水平微调PEG数量,以进一步提升增强效果与模型稳健性。



((a)original image;(b)denoised image;(c)noise)

图3 去噪结果可视化

Fig. 3 Visualization of Denoising Results

1.2 双域协同注意力去噪

为缓解低照图像中普遍存在的噪声干扰问题(Zhang等,2017),本文在分解前引入噪声抑制模块(Denoise Attention, DNA),以提升后续照明图与反射图建模的准确性与稳定性。DNA模块所关注的噪声主要表现为以信号无关的加性高斯噪声为主,并混合少量由光子统计波动引起的泊松噪声,这符合多数CMOS/CCD传感器在不同ISO设置下的典型

低光噪声特性。需要指出的是,本文并未对噪声分布形式进行显式建模,而是通过卷积特征提取与注意力加权机制,在特征域中自适应学习噪声与结构信号之间的统计差异(Foi等,2008)。该设计使DNA模块无需依赖具体传感器型号或ISO参数,即可在端到端训练过程中隐式适配不同光照条件下的混合噪声分布,从而实现稳健的低光去噪效果。该模块以浅层卷积和多层残差单元为基础,并结合去噪注

注意力机制对图像进行自适应去噪。输入灰度图像 $X \in \mathbb{R}^{1 \times H \times W}$ 首先通过卷积层提取初始特征:

$$F_0 = \delta(C(X)) \quad (1)$$

式中, F_0 表示由卷积层提取的初始特征图, $\delta(\cdot)$ 表示带泄露的修正线性单元激活函数 (leaky rectified linear unit, Leaky ReLU), C 表示卷积运算, X 表示输入的灰度低照度图像。随后, 特征图 F_0 依次通过 3 个带残差连接的卷积模块, 捕捉局部结构信息, 形成中间特征 F_r :

$$F_r = F_0 + \sum_{i=1}^3 R_i(F_i) \quad (2)$$

式中, F_r 表示经过多层残差卷积模块后的输出特征图, R_i 表示第 i 个残差卷积模块, F_i 表示第 i 个残差模块的输入特征。为了进一步增强网络对重要区域的关注能力, 引入了去噪注意力机制, 融合通道注意力 (Channel Attention) 与空间注意力 (Spatial Attention) 的联合注意力机制 (Woo 等, 2018)。通道注意力部分首先对输入特征进行全局平均池化, 再通过两层逐点卷积建模通道间依赖关系, 得到通道注意力权重:

$$M_c(F) = \sigma(W_2 \cdot \delta(W_1 \cdot P(F))) \quad (3)$$

式中, $M_c(F)$ 表示由通道注意力机制生成的特征图, F 表示输入特征图, σ 表示 Sigmoid 函数, W_1, W_2 为 1×1 卷积核, P 表示全局平均池化。空间注意力部分通过深度可分离卷积挖掘局部显著性区域, 其表达式为:

$$M_s(F) = \sigma(C_{1 \times 1}(C_{3 \times 3}^{dw}(F))) \quad (4)$$

式中, $M_s(F)$ 表示由空间注意力机制生成的特征图, $C_{3 \times 3}^{dw}$ 表示深度可分离卷积。两种注意力权重经逐点相乘后作用于原始特征图, 实现通道与空间维度的联合选择性增强:

$$F_a = F + F \cdot (M_c(F) \cdot M_s(F)) \quad (5)$$

式中, F_a 表示通道与空间维度融合后的特征图。最终, 输出特征经尾部卷积层和 Sigmoid 归一化后得到去噪后的图像:

$$\tilde{X} = \sigma(C(F_a)) \quad (6)$$

式中, 输出 $\tilde{X} \in [0, 1]^{1 \times H \times W}$ 将作为后续照明图和反射图估计的输入。图 3 中的噪声主要表现为低照条件下由传感器读出过程引入的高频随机扰动, 并在物体轮廓等高梯度区域更加显著。该噪声在真实成像中通常与泊松噪声混合, 但在本实验所示场景下, 其统计特性更接近加性高斯噪声, 因此 DNA 模块对该类

噪声的有效抑制为后续 Retinex 分解提供了更加稳定和可靠的输入基础。

1.3 混合残差与 U-Net 架构

本文方法用于表征图像全局与局部光照分布的照明图 L 应当具有平滑连续的空间特性, 同时能够覆盖强光、阴影等高动态区域。为此, 本文设计了多层残差卷积网络作为照明图估计子网络 L-Net, 并加入结构对称、感受野充足的 U-Net 架构辅佐多尺度光照信息的有效建模与融合, 同时 U-Net 也对后续 R-Net 网络建模反射图进行引导。

轻量级的网络 L-Net 由三部分组成: 头部、主体和尾部。头部包含一个反射填充层、一个卷积层和修正线性单元 (rectified linear unit, ReLU) 激活函数; 主体由三个残差块组成, 用于深层特征提取 (Pan 等, 2022); 尾部通过一个反射填充层和一个卷积层将通道数降为 1, 最后通过 *sigmoid* 函数得到归一化的光照图估计。U-Net 的输入为去噪后的图像 $X' \in \mathbb{R}^{1 \times H \times W}$, 网络首先通过一个双卷积模块提取初始特征, 随后依次经过三层下采样编码器逐步获取高层次语义信息, 最终再通过对称结构的三层上采样解码器逐级还原分辨率并融合低层特征, 实现光照图的重建:

$$\begin{aligned} F_0 &= C(X'), F_1 = D_1(F_0), \dots, F_3 = D_3(F_2), \\ F_4 &= Up_1(F_4, F_3), \dots, F_6 = Up_3(F_5, F_0), \\ L &= \sigma(C(F_6)) \end{aligned} \quad (7)$$

式中, X' 表示输入的特征图, D 表示下采样模块 (最大池化+双卷积), Up 表示上采样模块 (反卷积+跳跃连接+双卷积), F_i 表示采样后的特征图。每层跳跃连接允许低层的细节信息参与最终输出的恢复, 确保在估计光照图时保留边缘平滑性与区域一致性。为了进一步提升感受野并抑制边界伪影, 网络中所有卷积操作均配合 *Reflection Padding* 实现。此外, L-Net 的最终输出经过 *Sigmoid* 函数归一化至 $[0, 1]$ 区间, 满足 Retinex 理论对光照分量的定义。在 L-Net 的训练过程中, 为了确保分解的准确性和稳定性, 网络采用了多种损失函数来约束照明图 L 的学习过程, 如图 4 所示, 主要有 Retinex Loss 和 Guidance Loss。其中与照明图生成有关的损失函数共同构成了 Retinex Loss, 其中照明图本身的平滑性通过 Total Variation Loss 约束:

$$L_{TV}(L) = \|\nabla_x L\|_1 + \|\nabla_y L\|_1 \quad (8)$$

式中, L_{TV} 表示损失函数, L 表示光照分量, ∇_x 和 ∇_y 表示水平和垂直方向的梯度算子。
© 中国图象图形学报版权所有

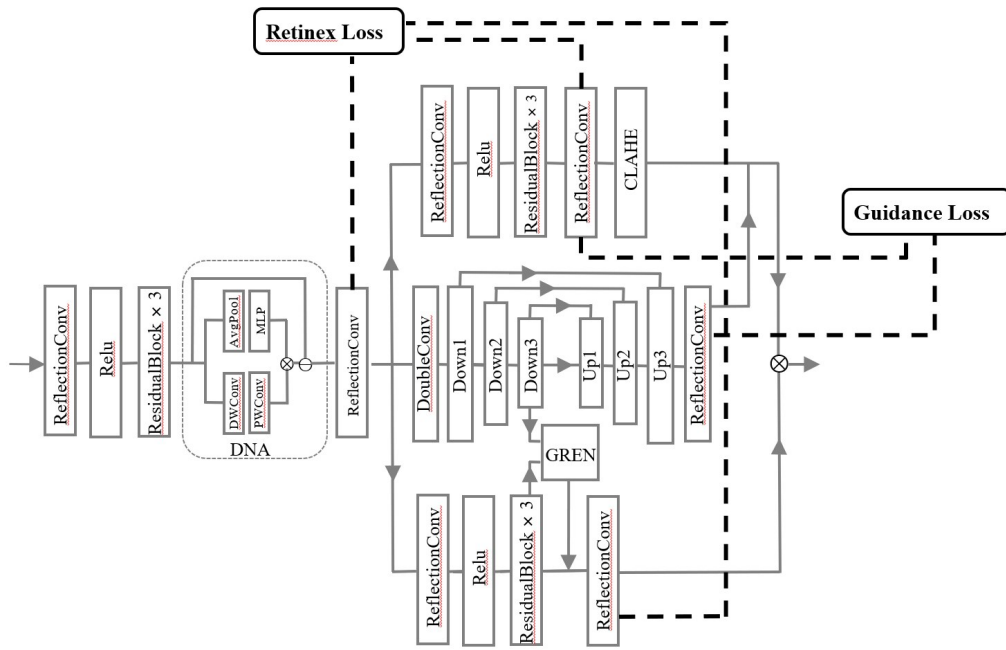


图4 损失函数设计示意图

Fig. 4 Overview of the Loss Function Design

表示图像在水平方向和垂直方向的梯度张量。也包括使用最大通道拟合来近似图像的明度:

$$L_i = MSE(L, \max(X')) \quad (9)$$

式中, L_i 表示损失函数, MSE 表示均方误差损失函数, L 表示预测光照图, X' 表示去噪后的输入图像, $\max()$ 表示沿通道维度取最大值以近似图像亮度。此外, 为保证 Retinex 分解满足理论上的乘积一致性, 网络在训练过程中引入了反射与照明的重构一致性约束。根据 Retinex 公式:

$$X = L \cdot R \Rightarrow R = \frac{X}{L} \quad (10)$$

式中, X 是原始特征图, L 是光照信息分量, R 是反射图分量。因此, 网络在分解结果满足物理模型的同时也对反射图 R 与反演值 $\frac{X}{L}$ 之间引入均方误差约束项:

$$L_r = \|L \cdot R - I\|_2^2 \quad (11)$$

$$L_d = \left\| R - \frac{I}{L + \epsilon} \right\|_2^2 \quad (12)$$

式中, L_d 表示损失函数, ϵ 为一个很小的常数 (如 1×10^{-6}), 用于防止除零异常。鉴于 U-Net 的下采样部分凭借多尺度特征提取机制, 能够高效捕捉到丰富的语义信息与精细的空间结构, 本文决定利用这一优势来为 R-Net 提供引导。为了确保 U-Net 的下采样部分能够充分吸收 L-Net 所蕴含的关键信息, 进而

为 R-Net 提供有效的指导, 引入了一个额外的损失函数 Guidance Loss。

$$L_g = \frac{1}{N} \sum_{i=1}^N |L_{u,i} - L_i| \quad (13)$$

式中, L_g 表示损失函数, N 表示参与损失计算的样本总数, L_u 为 U-Net 网络输出特征, L 为 L-Net 网络输出的特征。能够促使 U-Net 的输出尽可能地贴近 L-Net 的输出, 在模型训练过程中, 实现二者输出的高度一致性, 确保了信息在二者之间的顺畅传递, 为 R-Net 的优化训练奠定了坚实的基础。1.4 过曝感知的引导式反射估计

在本文分解框架中, 反射图含有图像中与光照无关的固有属性, 包括结构纹理和材质特征, 但分解时其中携带的不必要光照信息会引起局部区域过曝。为了抑制过曝现象并生成高质量的反射图估计, 本文设计了一个带有照明定向信息的引导式反射图增强模块 (GREN)。如图 5 所示, R-Net 接收由去噪子网络 (D-Net) 输出的图像 X' , 并通过多层残差卷积提取了 GREN 模块的初始输入特征之一 $F_R \in \mathbb{R}^{64 \times H \times W}$ 。

同时, 从与照明估计网络对齐的 U-Net 中提取深层语义特征 $F_L \in \mathbb{R}^{256 \times H/8 \times W/8}$, 通过两层 1×1 卷积将其通道数从 256 降至 64, 实现通道压缩, 并通过双线性插值对降维后的特征进行空间对齐, 确保与 F_R 在空

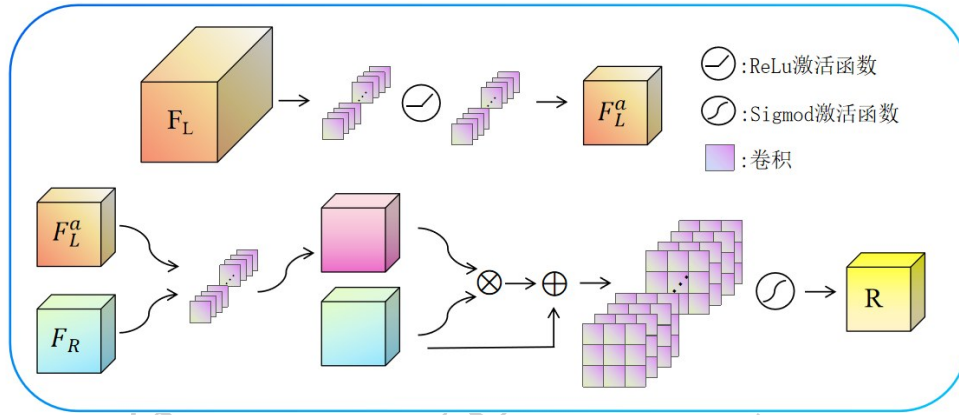


图5 GREN模块细节图

Fig. 5 Detailed Illustration of GREN

间维度的一致性:

$$F_L^d = \sigma(\delta(C_{1 \times 1}(F_L))) \quad (14)$$

$$F_L^a = I(F_L^d, F_R) \quad (15)$$

式中, F_L 表示 U-Net 中提取的深层语义特征, F_R 表示通过多层残差卷积提取的特征图, I 表示对特征图在空间维度进行对齐。随后将两个特征图在通道维度上拼接, 并通过一层 1×1 卷积进行压缩与融合, 得到引导增强后的特征表示:

$$F_f = C_{1 \times 1}([F_R; F_L^a]) \quad (16)$$

式中, F_L^a 表示 F_L^d 与 F_R 空间维度对齐的特征图。该引导融合机制能够显式地向 R-Net 提供照明上下文信息, 从而增强模型对边缘、纹理与阴影区域的判别

能力, 降低照明变化对反射估计的影响。融合特征 F_f 接着送尾部输出卷积进行编码, 进一步提取高阶结构信息后输出单通道反射图, 并使用 Sigmoid 激活函数归一化至 $[0, 1]$ 区间:

$$R = \sigma(C(F_f)) \quad (17)$$

式中, F_f 表示压缩融合后的特征图, R 表示 GREN 模块处理后的输出特征图。该网络设计通过引入照明引导特征 (Guo 等, 2025), 显著增强了反射图对结构与纹理的表达能力, 有效提升了光照与反射分量的分离质量, 为后续图像增强提供了稳定的反射信息支撑。



图6 WildNight-8K数据集典型样例图

Fig. 6 Typical Examples from the WildNight-8K Dataset

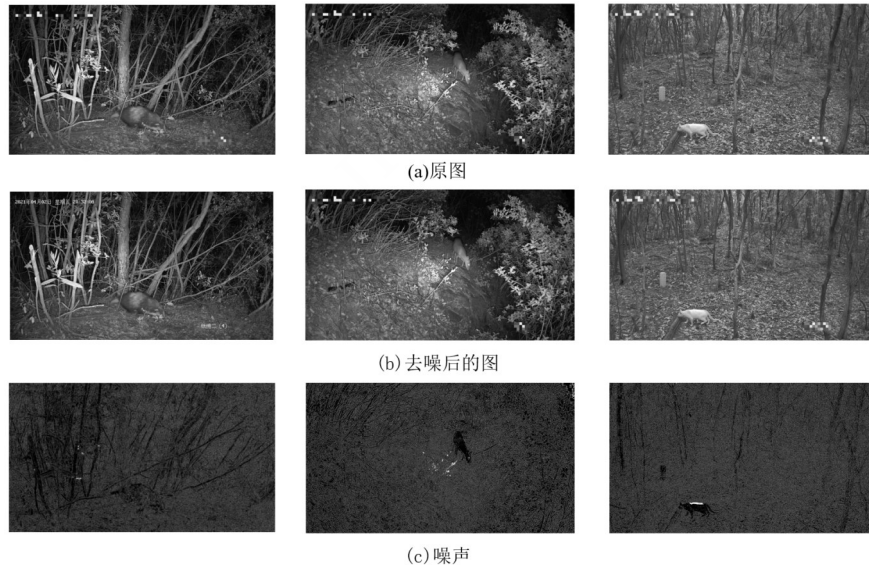
2 实验

2.1 数据集和实施细节 2.1.1 数据集介绍

本研究使用了3个数据集进行实验评估: 1个私有弱光图像数据集 WildNight-8K (Wildlife Night-time 8000-image dataset) 和2个公开数据集 LOL 和

DICM (Lee 等, 2013)。

其中, WildNight-8K, 包含 8000 张图像, 来源于自然动物保护区的 10 个监测摄像头, 分辨率为 3200×1800 。由于该数据集涵盖真实的野生动物夜间监控场景, 具有较强的实际代表性和复杂光照条件, 因此对本文研究具有重要意义。数据质量受限于监控设备性能和环境光照变化, 存在噪声与纹理



((a)original image;(b)low-light enhanced image)

图7 本文方法增强效果细节图

Fig. 7 Detailed image of the enhanced effect of this method

复杂性,且不包含对应的正常光照质量参考图,因此本文在该数据集上采用无监督指标进行评估。表2给出了详细统计信息,图6展示了部分典型样例。

此外,为验证本文方法在典型无参考弱光增强场景中的适用性,实验在DICM和LOL数据集上进行。DICM包含64张来自多种数码相机的真实低光图像,涵盖室内昏暗环境、室外夜景以及复杂光照条件,如背光和阴影区域;LOL数据集则提供500张配对的低光/正常光图像。两者的结合能够全面评估算法在不同实际场景下的鲁棒性与适应性。

表2 WildNight-8K数据集统计信息
Table 2 WildNight-8K Dataset Statistics

WildNight-8K数据集统计信息	
图片来源	浙江省东阳市东白山自然保护区
图片数量	8000
图片类别	夜间短波红外成像图
分辨率	3200×1800
拍摄时段	20:00 - 次日 5:00
拍摄场景	密林、灌木带、草地、藤蔓覆盖区等
摄像头总数	108
动物类别	果子狸、黄鹿、猪獾、白鹇、野猪等
亮度平均值	74.75
亮度最小值	64.82
亮度最大值	93.02

2.1.2 实施细节

本文使用自适应矩估计优化器(adaptive moment estimation, Adam),设置 $\beta_1=0.9$, $\beta_2=0.999$,初始学习率设置为 1×10^{-4} 。训练过程中采用多步学习率衰减策略,每训练100个epoch将学习率衰减为原来的0.5倍。训练批次大小设为4,总共训练300个epoch。训练使用单张NVIDIA GPU 4060Ti(启用了CUDA加速)。模型的训练过程包括4种损失函数的联合优化,为防止训练过程中的不稳定情况,加入了NaN检测逻辑,并在每次反向传播后应用了最大范数为1.0的梯度裁剪,以稳定模型更新过程。

2.2 评价指标

为了客观且全面地评估所提出图像增强算法在弱光条件下的性能表现,本文选用了五种广泛应用且具有代表性的图像质量评价指标,包括三种传统无参考指标:自然图像质量评估指标(natural image quality evaluator, NIQE)(Mittal等,2013)、无参考图像空间质量评估器(blind/referenceless image spatial quality evaluator, BRISQUE)(Mittal等,2012)和基于深度残差网络DRN指标(deep residual network, DRN)(Zhang等,2017),以及两种感知导向指标:感知指数(perceptual index, PI)(Blau等,2018)和学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)(Zhang等,2018)。其中,NIQE通过衡量图像统计特征与自然图像模型的差异来评估

自然度, BRISQUE 侧重提取局部空间特征以评估失真程度, DRN 利用深度学习模型模拟人类主观感知以更精准地评价视觉质量, 而 PI 通过无参考指标提供更符合人类视觉感知的整体评分, LPIPS 基于深度特征空间的感知相似性度量, 可以有效反映增强后图像在结构与纹理上的变化。通过这五种指标的综合分析, 能够更加全面地反映算法在亮度改善、细节保留以及感知自然度等方面的性能表现。

2.3 对比实验

本研究算法在真实弱光场景下的增强效果如图 7 所示, 在有效提升图像亮度的同时, 避免了过曝现象的出现, 并显著增强了图像的可见性。增强后的图像在保持自然亮度分布的基础上, 展现出更丰富的纹理和细节信息, 且未引入明显的噪声或伪影, 整体视觉质量得到显著改善。

为了验证本文方法的有效性和鲁棒性, 分别在 WildNight-8K、DICM 和 LOL 数据集上与 9 种代表性弱光图像增强方法进行了定量与定性对比分析。所选方法涵盖了经典的传统算法与近年来提出的主流深度学习模型, 包括 ZERO-DCE (Guo 等, 2020)、RetinexNet (Dong 等, 2022)、RUAS (Li 等, 2021)、EnlightenGAN (Jiang 等, 2021)、SCI (Ma 等, 2022)、RetinexFormer (Cai 等, 2023)、LightenDiffusion (Jiang 等, 2024)、FLOL (Benito 等, 2025) 以及 CIDNet (Yan 等,

2025)。本文采用 NIQE、BRISQUE 和 DRN 三种无参考图像质量评价指标, 对各方法增强结果的客观性能进行评估, 结果见表 3 和表 4。

为了验证本文方法的有效性和鲁棒性, 分别在 WildNight-8K 数据集、DICM 数据集和 LOL 数据集上与 9 种代表性弱光图像增强方法进行了定量与定性对比分析。所选方法涵盖了经典的传统算法与近年来提出的主流深度学习模型, 包括 ZERO-DCE (Guo 等, 2020)、RetinexNet (Dong 等, 2022)、RUAS (Li 等, 2021)、EnlightenGAN (Jiang 等, 2021)、SCI (Ma 等, 2022)、RetinexFormer (Cai 等, 2023)、LightenDiffusion (Jiang 等, 2024)、FLOL (Benito 等, 2025) 以及 CIDNet (Yan 等, 2025)。本文采用 NIQE、BRISQUE 和 DRN 三种无参考图像质量评价指标, 对各方法增强结果的客观性能进行评估, 结果见表 3 和表 4。

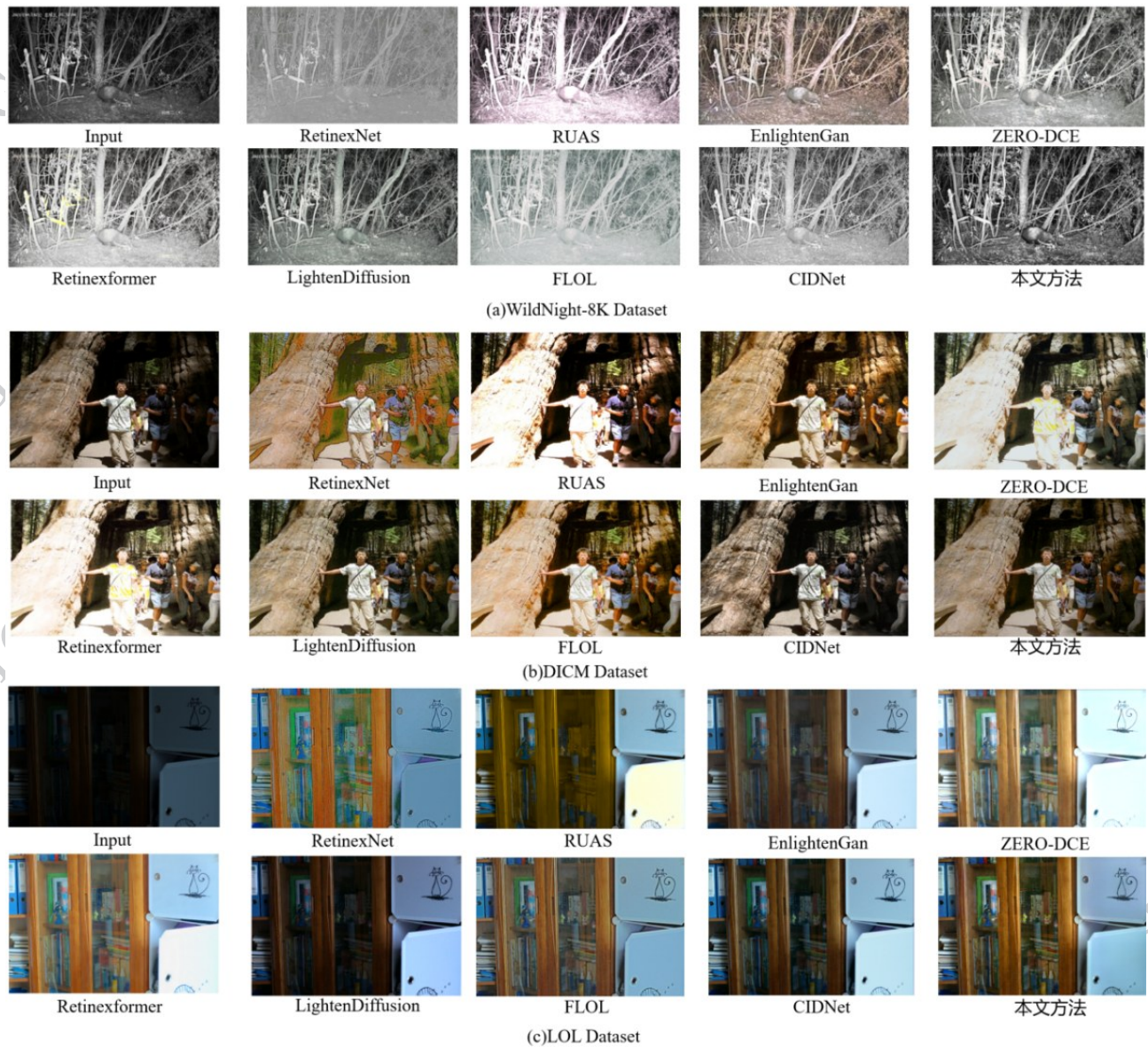
同时, 图 8 在 WildNight-8K、DICM 和 LOL 数据集上给出了各方法的视觉对比。在 WildNight-8K 数据集上原始图像整体黯淡, 仅零星区域出现亮斑, RetinexNet 虽抬升局部暗部亮度, 但整体偏暗且伴随明显糊化与噪声放大; RUAS、ZERO-DCE、Retinexformer 与 FLOL 在增亮方面表现积极, 但易引入大面积过曝。EnlightenGAN、LightenDiffusion 及 FLOL 虽抑制了高光溢出, 却整体色调偏黄偏绿; CIDNet 在降噪与亮度提升上较为均衡, 然而纹理细节仍显粗

表 3 不同模型在 LOL 数据集上的对比实验

Table 3 Comparative experiments of different models on LOL datasets

模型	LOL							
	类型	Params (M)	FLOPs (G)	NIQE↓	BRISQUE↓	DRN↓	LPIPS↓	PI↓
RetinexNet (Dong 等, 2022)	有监督	0.84	587.47	32.9662	25.6694	0.4128	0.3979	19.1942
RUAS (Li 等, 2021)	有监督	0.003	0.83	27.2001	20.3618	0.7808	0.2707	14.8577
SCI (Ma 等, 2022)	有监督	0.0035	0.69	30.1886	22.9022	0.6014	0.3497	18.4174
Retinexformer (Cai 等, 2023)	有监督	1.61	15.57	26.9033	14.7295	0.3583	0.2821	15.0211
CIDNet (Yan 等, 2025)	有监督	1.88	7.57	28.5411	19.9677	0.4005	0.2979	13.8376
ZERO-DCE (Guo 等, 2020)	无监督	0.075	4.83	27.9745	20.2504	0.5982	0.3305	15.5021
EnlightenGAN (Jiang 等, 2021)	无监督	114.35	61.01	28.1024	16.2268	0.6201	0.3097	14.9981
LightenDiffusion (Jiang 等, 2024)	无监督	27.81	59.32	27.2058	15.0859	0.5256	0.3383	14.0445
FLOL (Benito 等, 2025)	无监督	0.81	8.36	28.3371	21.1203	0.3422	0.2963	15.2402
本文方法	无监督	2.64	17.32	26.2812	14.5511	0.3376	0.2611	13.7917

注: 加粗字体表示最优结果; 箭头向下表示负向指标, 数值越低代表表现越优。



((a) WildNight-8K Dataset; (b) DICM Dataset; (c) LOL Dataset)

图 8 不同模型在图像增强任务中的可视化结果对比

Fig. 8 Visual Comparison of Enhancement Results Across Different Models

糙。同时在 DICM 数据集上的实验结果表明, RetinexNet 在提升图像亮度的同时引入了较多噪声; RUAS、ZERO-DCE 和 Retinexformer 均存在明显的过曝光现象; EnlightenGAN 和 FLOL 虽能有效增强亮度,但整体色调偏黄,存在色差问题; LightenDiffusion 和 CIDNet 虽未出现上述问题,但与本文方法相比,对图像中后段背景的亮度提升效果有限,且在树木等纹理细节的还原上不够自然。最后在 LOL 数据集上,各方法在亮度恢复与色彩一致性方面的差异明显, LOL 场景为室内弱光日常环境, RetinexNet 在该数据集上依旧存在分解不稳定的问题,增强后图

像出现明显的颜色偏移与伪彩现象,且暗部噪声被进一步放大; RUAS 与 ZERO-DCE 虽能显著提升整体亮度,但伴随局部区域过曝,尤其在前景高反射物体处细节丢失严重; Retinexformer 在亮度提升上较为充分,但仍存在亮区泛白和柜子颜色失真的问题。FLOL 在一定程度上缓解了过曝现象,但增强结果整体偏黄偏暖,色彩还原准确性不足; EnlightenGAN、LightenDiffusion 和 CIDNet 在色彩稳定性方面表现较好,噪声控制相对理想,但对深暗区域(书本部分)的提亮幅度有限,导致整体观感仍偏暗,部分低对比纹理难以辨识。综上所述,本文方法在显著提亮全局

表 4 不同模型在 WildNight-8K 和 DICM 数据集上的对比实验

Table 4 Comparative experiments of different models on WildNight-8K and DICM datasets

模型	WildNight-8K					DICM			
	类型	NIQE↓	BRISQUE↓	DRN↓	PI↓	NIQE↓	BRISQUE↓	DRN↓	PI↓
RetinexNet(Dong等,2022)	有监督	47.7723	33.0181	0.1986	27.5802	32.8056	30.3956	0.6652	21.1579
RUAS(Li等,2021)	有监督	25.3308	37.9463	1.1967	21.4092	28.0808	24.0853	1.2288	18.9102
SCI(Ma等,2022)	有监督	39.4766	31.3406	0.9984	24.257	35.7977	29.3843	0.6653	22.4673
Retinexformer(Cai等,2023)	有监督	23.5858	28.4914	0.204	15.9012	27.3817	16.1676	0.5749	18.6652
CIDNet(Yan等,2025)	有监督	26.0875	33.4579	0.2463	16.1024	28.4569	24.8822	0.5967	18.7355
ZERO-DCE(Guo等,2020)	无监督	25.6587	35.2875	0.5524	21.9802	30.2217	24.2588	0.6995	20.1452
EnlightenGAN(Jiang等,2021)	无监督	24.1287	44.5487	0.3565	21.5433	28.7177	18.9548	0.7803	19.3542
LightenDiffusion(Jiang等,2024)	无监督	24.0353	31.5743	0.1896	16.5801	28.1835	26.5121	0.6867	18.8823
FLOL(Benito等,2025)	无监督	26.4565	31.1065	0.6513	17.2581	29.4709	23.2186	0.5788	19.4211
本文方法	无监督	22.2610	27.1911	0.1814	15.5665	27.2728	15.1049	0.5642	18.1875

注:加粗字体表示最优结果;箭头向下表示负向指标,数值越低代表表现越优。

的同时,精细还原了远景植被、近景纹理,有效抑制过曝与噪声,亮度自然、细节丰富,视觉效果最接近真实光照场景(Dong等,2025)。

2.4 消融实验

为了验证本文提出的三项关键模块——伪曝光图像生成(PEG)、去噪注意力机制(DNA)以及光照图引导的反射图增强(GREN)在模型中的有效性。如表5所示,以去除全部模块的模型为基准,在WildNight-8K上依次加入单一模块重新训练:加入PEG后反射细节明显,NIQE与DRN指标均好转,说明该模块可以引导网络学习更有效的结构细节,有助于重建低光条件下模糊或缺失的反射信息;加入DNA后噪声与过曝明显减少,NIQE和DRN指标分别下降11.6%和15.4%,验证了该注意力机制对噪声抑制的有效性;加入GREN则促进了光照与反射的平衡,图像结构和细节还原方面变得自然,过曝区域显著减少。相比之下,加入全部模块的完整模型在所有无参考指标上均保持最优,增强图像在亮度、清晰度和自然感方面表现最优,充分证明各模块协同增益的重要性。

为了进一步系统性验证本文提出的伪曝光图像生成(PEG)、去噪注意力机制(DNA)以及光照图引导的反射图增强(GREN)在整体框架中的独立贡献与协同作用,本文开展了更加完善的消融实验,如表

表 5 增量消融实验表

Table 5 Incremental Ablation Study Table

Baseline	PEG	DNA	GREN	NIQE↓	DRN↓
√				27.4093	0.3109
√	√			25.3557	0.2822
√		√		24.2370	0.2631
√			√	24.2909	0.2620
√	√	√	√	22.2610	0.1814

注:加粗字体表示最优结果;√表示采用相应部分;箭头向下表示负向指标,数值越低代表表现越优。

表 6 消融实验表

Table 6 Ablation Study Table

模型	PEG	DNA	GREN	NIQE↓	DRN↓
w/o PEG		√	√	24.2107	0.2402
w/o DNA	√		√	23.9627	0.2355
w/o GREN	√	√		23.8566	0.2258
完整模型	√	√	√	22.2610	0.1814

注:加粗字体表示最优结果;√表示采用相应部分;箭头向下表示负向指标,数值越低代表表现越优。

6所示,以包含全部模块的完整模型为参照,分别构建w/o PEG、w/o DNA和w/o GREN的对比模型。结果表明,去除PEG后,模型在暗区结构恢复能力明显下降,反射细节模糊,NIQE与DRN分别上升至

24. 2107 和 0. 2402, 说明伪曝光样本在引导网络学习稳定反射结构方面起到了关键作用; 去除 DNA 后, 噪声抑制能力显著减弱, 高亮区域出现明显过曝, NIQE 与 DRN 均明显劣化, 验证了去噪注意力机制在抑制噪声放大和高光失真中的核心作用; 去除 GREN 则导致光照与反射分离不充分, 图像整体结构虽保持, 但局部亮区过曝、纹理不自然, NIQE 和 DRN 分别为 23. 8566 和 0. 2258。相比之下, 完整模型在两项无参考指标上均取得最优结果, 充分说明三者亮度提升、噪声抑制与结构恢复之间形成了有效的协同增益。

表 7 复杂度消融分析表

Table 7 Ablation Study on Model Efficiency (Params/FLOPs)

模型	Params(M)	FLOPs(G)
w/o NDA	2.42	13.67
w/o GREN	2.59	13.54
完整模型	2.64	17.32

注: 加粗字体表示最优结果。

其次, 为排除性能提升仅来源于模型规模扩大的可能性, 本文进一步对 DNA 与 GREN 模块引入前后的参数量与计算复杂度进行了对比分析, PEG 属于数据层面的伪曝光生成策略, 而非可学习的网络模块, 在推理阶段不会引入额外的可训练参数或计算开销。结果如表 7 显示, 相比 w/o DNA 与 w/o GREN 的模型, 完整模型参数量仅增加至 2. 64M, FLOPs 为 17. 32G, 增幅有限, 却带来了显著的 NIQE 与 DRN 改善, 说明本文方法在计算开销可控的前提下实现了更高的增强质量, 具有良好的效率 - 性能平衡。

最后, 本文进一步对 backbone 中 U-Net 的编码器 - 解码器结构配置进行了系统对比, 不仅考察了不同网络深度的对称结构, 还分析了非对称结构对低光增强性能的影响。实验结果如表 8 所示, 当 Encoder/Decoder 均为 2 层时, 网络表达能力受限, 难以充分建模复杂低光场景中光照分布与噪声耦合关系, 整体增强效果较弱; 随着网络加深至对称的 3 层结构, 模型在 NIQE 和 DRN 指标上取得显著提升 (22. 2610/0. 1814), 在亮度恢复、噪声抑制与细节保持之间达到最佳平衡; 进一步加深至对称的 4 层结

表 8 U-Net 网络深度与编码/解码对称性的消融实验

Table 8 Ablation Study on U-Net Depth and Encoder/Decoder Symmetry

U-Net 结构	NIQE ↓	DRN ↓	Params (M)	FLOPs(G)
E2 - D2	24.7421	0.2366	1.46	16.82
E3 - D3	22.2610	0.1814	2.64	17.32
E4 - D4	23.0211	0.1899	8.51	18.13
E3 - D4	23.1801	0.1977	2.65	17.56
E4 - D3	23.5789	0.2143	8.47	17.54

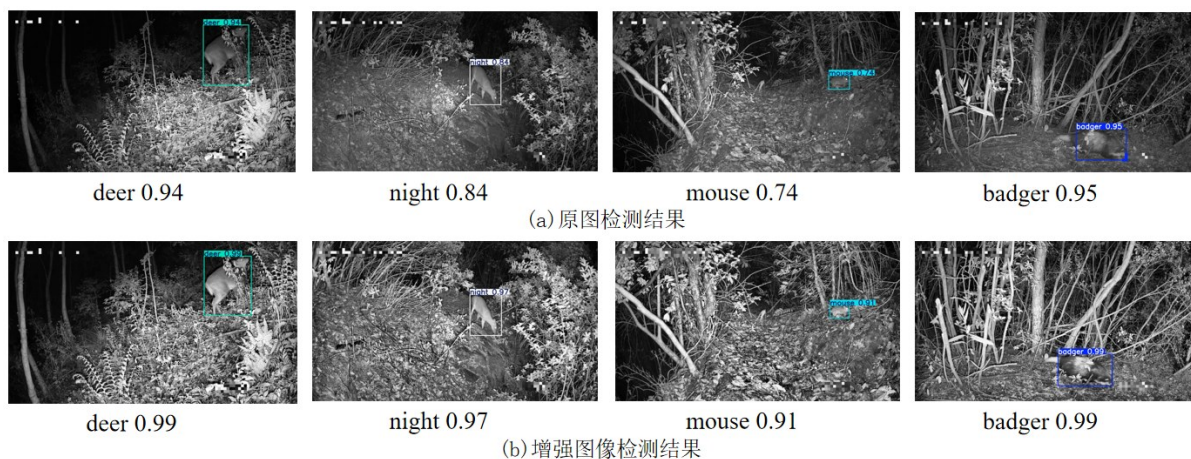
注: 加粗字体表示最优结果; 箭头向下表示负向指标, 数值越低代表表现越优。

构时, 尽管参数量和计算量明显增加, 但性能提升趋于饱和甚至出现回退, 说明过深网络容易引入冗余特征建模, 并可能干扰光照和反射分解的稳定性。与此同时, 本文还对非对称 U-Net 结构——Encoder4 层 - Decoder3 层和 Encoder3 层 - Decoder4 层进行了对比实验。实验结果表明, 编码器与解码器层级匹配对于多尺度光照信息的稳定传递至关重要, 对称的 U-Net 结构能够在下采样与上采样过程中保持光照语义与空间尺度的一致性, 使不同尺度的光照特征得以逐级、有效地恢复; 而非对称结构或过深网络易破坏这种尺度对齐, 导致光照建模冗余或局部失真。

综上所述, 多角度消融实验从模块有效性、复杂度控制以及网络结构设计等方面全面验证了本文方法的合理性与鲁棒性, 进一步说明所提出的 PEG、DNA 与 GREN 模块在低光增强任务中各司其职、相互协同, 共同促成了整体性能的显著提升。

2.5 基于下游任务的评估

为了进一步验证所提出弱光图像增强方法在真实应用场景中的实用价值, 开展了基于任务的评估实验 (Task-Driven Evaluation), 将增强后的图像应用于下游的动物目标检测任务, 考察其对目标识别性能的影响。具体而言, 选用当前性能领先的检测模型 YOLOv8 (Jocher, 2023), 并以 WildNight-8K 数据集中经不同增强方法处理后的图像作为训练样本, 构建动物检测任务数据集。所有检测模型均在统一配置下训练, 唯一变量为图像增强策略。本文采用主流的目标检测评价指标 mAP (mean Average Precision) 在测试集上对检测性能进行评估。实验结果



((a)detection results on original images;(b)detection results on enhanced images)

图9 YOLOv8 检测结果的可视化对比

Fig. 9 Visual Comparison of YOLOv8 Detection Results

如表9所示,使用本文方法增强的图像作为训练数据,YOLOv8获得了显著更高的检测精度,在各项性能指标上均优于原始弱光图像及其他增强方法。本文增强图像在检测中的可视化示例如图9所示,进一步从任务导向的角度验证了本方法在提升图像可见性、结构可读性与真实感方面的有效性,展现出其在实际应用中的强大潜力与泛化能力。

表9 不同增强模型在YOLOv8检测中的性能对比

Table 9 Comparison of YOLOv8 Detection Performance on Images Enhanced by Different Models

模型	mAP50-95(YOLOv8)
原图	0.827
RetinexNet(Dong等,2022)	0.838
RUAS(Li等,2021)	0.863
EnlightenGAN(Jiang等,2021)	0.847
SCI(Ma等,2022)	0.849
Retinexformer(Cai等,2023)	0.841
Retinexmamba(Bai等,2024)	0.859
LightenDiffusion(Jiang等,2024)	0.842
FLOL(Benito等,2025)	0.837
Cidnet(Yan等,2025)	0.843
本文方法	0.873

注:加粗字体表示最优结果。

3 结论

针对弱光条件下图像易出现亮度不足、噪声放大、局部过曝及结构细节丢失等问题,本文提出了一种基于光照分量引导反射分量估计的零监督弱光图像增强方法。在无需成对正常光照数据的前提下,引入伪曝光生成策略构建多亮度训练样本,结合去噪注意力机制与光照引导的反射增强模块,构建了一个弱光增强框架,实现了光照与反射分量的稳定分解与高质量重建。在WildNight-8K、DICM和LOL三个弱光数据集上的实验结果表明,本文方法在NIQE、PI和DRN等多项评价指标上整体优于RetinexNet、RUAS和ZERO-DCE等代表性方法,能够在提升图像亮度的同时有效抑制过曝与噪声放大,并较好地保持结构与纹理细节。进一步的YOLOv8目标检测实验显示,基于本文方法增强的图像在mAP@50-95指标上较原始弱光图像取得明显提升,验证了该方法对下游感知任务的有效促进作用。尽管本文方法在增强质量和鲁棒性方面取得了较好效果,但仍存在一定局限性,如模型结构相对复杂、计算开销较高,以及在极端噪声或强局部光源条件下反射与光照分离仍不完全。未来将重点开展模型轻量化设计,进一步提升复杂弱光场景下的稳定性,并拓展其在更多下游视觉任务和实际应用中的适用性。

参考文献

- Bai J, Yin Y, He Q, Zhang W and Liu F. 2024. RetinexMamba: Retinex-based Mamba for low-light image enhancement [EB/OL]. [2024-06-01].
<https://arxiv.org/abs/2405.03349>
- Benito J C, Feijoo D, Garcia A, Alvarez M and Santos R. 2025. FLOL: Fast baselines for real-world low-light enhancement [EB/OL]. [2025-01-17].
<https://arxiv.org/abs/2501.09718>
- Blau Y, Mechrez R, Timofte R, Michaeli T and Zelnik-Manor L. 2018. The 2018 PIRM challenge on perceptual image super-resolution// Proceedings of the 2018 PIRM Workshop on Perceptual Image Restoration and Manipulation (associated with ECCV 2018). Munich, Germany: Springer: 1 - 12 [DOI:10.1007/978-3-030-11015-4_1]
- Cai J, Gu S and Zhang L. 2018. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4): 2049-2062 [DOI:10.1109/TIP.2018.2794218]
- Cai Y, Bian H, Lin J, Zhao K and Chen Y. 2023. Retinexformer: One-stage Retinex-based transformer for low-light image enhancement// Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE: 12504 - 12513 [DOI:10.1109/ICCV51070.2023.01149]
- Chen L, Liu Z, Fan X, Wang J and Zhao H. 2024. Unsupervised night-time image enhancement with spectral sensitivity awareness. *IEEE Transactions on Image Processing*, 33: 2122-2135 [DOI:10.1109/TIP.2024.3367892]
- Dong J, Li Y, Wang H, Xu F and Zhao W. 2022. Image enhancement using deep convolutional neural networks for night-time driving. *IEEE Transactions on Vehicular Technology*, 71(3): 3160-3171 [DOI:10.1109/TVT.2022.3145678]
- Dong W, Min Y, Zhou H and Chen J. 2025. Towards scale-aware low-light enhancement via structure-guided transformer design// Proceedings of the CVPR Workshops (NTIRE 2025). New Orleans, USA: IEEE: 1460-1470
- Foi A, Trimeche M, Katkovnik V and Egiazarian K. 2008. Practical Poissonian - Gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10): 1737 - 1754
- Guo C, Li C, Guo J, Zhou H and Wang M. 2020. Zero-reference deep curve estimation for low-light image enhancement// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 1780-1789 [DOI:10.1109/CVPR42600.2020.00182]
- Guo X, Li Y and Ling H. 2017. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2): 982-993 [DOI:10.1109/TIP.2016.2639450]
- Guo X, Tian Z, Wang Y, Li S, Jiang Y, Du S and Gao Y. 2025. ERet-inex: Event camera meets Retinex theory for low-light image enhancement// Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). Philadelphia, USA: IEEE: 3759 - 3765
- Ho J, Jain A and Abbeel P. 2020. Denoising diffusion probabilistic models// Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS). Virtual Conference: Curran Associates, Inc.: 6840 - 6851
- Jiang H, Luo A, Liu X, Han S and Liu S. 2024. LightenDiffusion: Unsupervised low-light image enhancement with latent-Retinex diffusion models// Proceedings of the European Conference on Computer Vision (ECCV). Glasgow, UK: Springer: 161 - 179 [DOI:10.1007/978-3-031-73195-2_10]
- Jiang Y, Gong X, Liu D, Cheng Y and Qiao Y. 2021. EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30: 2340-2349 [DOI:10.1109/TIP.2021.3051462]
- Land E H and McCann J J. 1971. Lightness and Retinex theory. *Journal of the Optical Society of America*, 61(1): 1-11 [DOI:10.1364/JOSA.61.000001]
- Lee C, Lee C and Kim C S. 2013. Contrast enhancement based on layered difference representation of 2D histograms. *IEEE Transactions on Image Processing*, 22(12): 5372 - 5385 [DOI:10.1109/TIP.2013.2282608]
- Li Z, Yang S, Yang H, Tang X, Wu F and Xu F. 2025. BIAWDiff: Enhancing low-light images with bio-inspired attention and wavelet diffusion// Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Milan, Italy: IEEE: 1 - 5 [DOI:10.1109/icassp49660.2025.10888793]
- Lin Y, Fu Z, Wen K, Ye T, Chen S, Meng G, Wang Y, Huang Y, Tu X and Ding X. 2025. DPLUT: Unsupervised low-light image enhancement with lookup tables and diffusion priors// Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). Spokane, USA: AAAI Press: 5316 - 5324
- Liu T, Zhang Y and Chen H. 2023. A review of image enhancement techniques for low-light and night vision applications in autonomous driving. *Sensors*, 23(7): 1714 [DOI:10.3390/s23071714]
- Ma L, Ma T, Liu R, Wang S and Zhao Q. 2022. Toward fast, flexible, and robust low-light image enhancement// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 5637-5646
- Mittal A, Moorthy A K and Bovik A C. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12): 4695-4708 [DOI:10.1109/TIP.2012.2214050]
- Mittal A, Soundararajan R and Bovik A C. 2013. Making a 'completely blind' image quality analyzer. *IEEE Signal Processing Letters*, 20(3): 209-212 [DOI:10.1109/LSP.2012.2227726]
- Pan X Y, Wei M, Wang H and Jia F Z. 2022. A multi-scale fusion

- residual encoder-decoder approach for low illumination image enhancement. *Journal of Computer-Aided Design & Computer Graphics*, 34(1): 104-112 (潘晓英, 魏苗, 王昊, 贾丰竹. 2022. 多尺度融合残差编解码器的低照度图像增强方法. *计算机辅助设计与图形学报*, 34(1): 104-112) [DOI:10.3724/SP.J.1089.2022.18833]
- Somepalli G, Singla A, Golatkar A, Bubeck S and Ge Y. 2023. DiffusionRetinex: Retinex-inspired low-light image enhancement via stochastic diffusion[EB/OL]. [2024-06-01]. <https://arxiv.org/abs/2303.09567>
- Sun W Q, Peng C Y and Zhang X J. 2025. A low-light image enhancement network via fused multi-scale dilated convolutions and coordinate grouping[J/OL]. *Journal of Image and Graphics*, 1-17 (孙婉倩, 彭春燕, 张效娟. 2025. 融合多重空洞卷积与坐标分组的暗光图像增强网络[J/OL]. *中国图象图形学报*, 1-17 [DOI: 10.11834/jig.250177])
- Wang Y, Liu Z, Liu J, Xu S and Liu S. 2023. Low-light image enhancement with illumination-aware gamma correction and complete image modelling network//*Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France: IEEE: 13128-13137
- Wei C, Wang W, Yang W, Liu J and Guo Y. 2018. Deep Retinex decomposition for low-light enhancement//*Proceedings of the British Machine Vision Conference (BMVC)*. Newcastle, UK: BMVA Press: 122
- Woo S, Park J, Lee J Y and Kweon I S. 2018. CBAM: Convolutional block attention module//*Proceedings of the European Conference on Computer Vision (ECCV)*. Munich, Germany: Springer: 3-19
- Xu L and Wang X. 2023. Contrast enhancement techniques for weakly lit driving scenes using multi-scale networks. *IEEE Access*, 11: 31244-31256 [DOI:10.1109/ACCESS.2023.3245678]
- Yan Q, Feng Y, Zhang C, Pang G, Shi K, Wu P, Dong W, Sun J and Zhang Y. 2025. HVI: A new color space for low-light image enhancement//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, TN, USA: IEEE: 5678 - 5687
- Zhang K, Zuo W, Chen Y, Meng D and Zhang L. 2017. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7): 3142-3155 [DOI:10.1109/TIP.2017.2699523]

作者简介

包晓安, 1973年生, 男, 教授, 主要研究方向为计算机视觉。

E-mail: baoxiaoan@zstu.edu.cn

涂小妹, 1995年生, 女, 讲师, 主要研究方向为图像处理、模式识别。E-mail: txm_95@163.com

陈奕江, 2000年生, 男, 硕士研究生, 主要研究方向为图像处理与机器视觉。E-mail: 1281520471@qq.com。

张娜, 1977年生, 女, 教授, 主要研究方向为图像识别。E-mail: zhangna@zstu.edu.cn。

胡天缤, 1998年生, 女, 博士研究生, 主要研究方向为人工智能。E-mail: 230222010006@hhu.edu.cn。

许铭洋, 2000年生, 男, 硕士研究生, 主要研究方向为图像处理与机器视觉。E-mail: 1791595405@qq.com。